# The Cusum Test for Parameter Change in Time Series Models

SANGYEOL LEE, JEONGCHEOL HA and OKYOUNG NA

*Seoul National University*

SEONGRYONG NA

*Yonsei University*

ABSTRACT. In this paper, we consider the problem of testing for parameter changes in time series models based on a cusum test. Although the test procedure is well established for the mean and variance in time series models, a general parameter case has not been discussed in the literature. Therefore, here we develop a cusum test for parameter change in a more general framework. As an example, we consider the change of the parameters in a random coeefficient autoregressive (1) model and that of the autocovariances of a linear process. Simulation results are reported for illustration.

Key words: autocovariance function, cusum test, invariance principle, linear process, martingale difference, RCA model, testing for parameter change, weak convergence

## 1. Introduction

Since the paper of Page (1955), the problem of testing for a parameter change has been an important issue among statisticians. Originally, the problem began with i.i.d samples; see Hinkley (1971), Brown *et al.* (1975), Zacks (1983), Csörgő & Horváth (1988), Krishnaiah & Miao (1988), Inclán & Tiao (1994), and it moved naturally into the time series context as economic time series often exhibit prominent evidence for structural change in the underlying model; see, for example, Wichern *et al.* (1976), Picard (1985), Krämer *et al.* (1988), Tang & MacNeil (1993), Kim *et al.* (2000), Lee & Park (2001), and the papers cited therein. If the random observations are i.i.d and follow a specific parametric model, one may consider utilizing a likelihood ratio method as in Csörgő & Horváth (1997). However, the method is no longer applicable if the underlying distribution is completely unknown. In such a case, a nonparametric approach should be considered as an alternative. From this viewpoint, here we pay attention to the cusum method for testing for parameter change.

The cusum method is easy to handle and useful for detecting the locations of change points as seen in Inclán & Tiao (1994). In particular, it has been utilized for testing for a change of mean, variance and distribution function (cf. Bai, 1994). A convenience of the method lies in the fact that the sample mean, variance and distribution function are all expressed as the sum of i.i.d random variables, and the convergence result of the cusum test statistic is easily obtained. In fact, Nyblom (1989) considered a sort of cusum method to handle the change point problem for parameters other than the mean and variance. However, the test procedure assumes that the underlying distribution of observations belongs to a known distribution family, and the test statistic is based on estimators relying on the underlying distribution. Unlike in his approach, here we pursue a cusum test, which is totally free from assumptions about the underlying distribution.

In fact, our cusum test can be constructed by imitating the one for a mean change in the i.i.d sample. Conventionally, the estimators of a target parameter after normalization are expressed as the sum of the average of i.i.d random variables and a negligible term. The basic idea is then

to view the change problem for the parameter as the one for the expected value of the random variables in that expression as a change of parameter would affect the expected value. Following this idea, one can easily construct the cusum test statistic. The details are presented in section 2.

The rest of this paper is organized as follows. In section 2, we present how the cusum test is constructed in a general framework. In section 3, we apply the cusum method to the problem of testing parameter constancy in random coefficient autoregressive (RCA) models based on a (conditional) least-squares estimator (LSE). Also, in section 4, we consider the problem of testing for an autocovariance change in infinite order moving average processes based on the sample autocovariance functions. Section 5 reports simulation results for the cusum tests regarding the models discussed in sections 3 and 4.

## 2. Cusum test

Here we explain how the cusum test is constructed. As an illustration, we consider the test for a mean change in an i.i.d sample based on the following process

$$
\begin{aligned}
U_n(s) &:= \frac{1}{\sqrt{n}\sigma}\left(\sum_{t=1}^{[ns]} x_t - \left(\frac{[ns]}{n}\right)\sum_{t=1}^{n} x_t\right) \\
&= \frac{[ns]}{\sqrt{n}\sigma}(\hat{\mu}_{[ns]} - \hat{\mu}_n), \quad 0 \le s \le 1,
\end{aligned} \tag{1}
$$

where $x_1, \ldots, x_n$ are i.i.d with mean $\mu$ and variance $\sigma^2$, and $\hat{\mu}_n = n^{-1}\sum_{t=1}^{n} x_t$. It is well known that $\{U_n\}$ converges weakly to a standard Brownian bridge, and a test is performed based on the convergence result. Similar reasoning can be adopted for the more general case.

Suppose that one is interested in testing for a change of $\theta$ based on a consistent estimator $\hat{\theta}_n$. As with the maximum likelihood estimator (MLE), usually $\hat{\theta}_n$ can be written as

$$
\hat{\theta}_n - \theta = n^{-1}\sum_{t=1}^{n} l_t + o_P\left(\frac{1}{\sqrt{n}}\right)
$$

(cf. Durbin, 1973), where $l_t := l_t(\theta)$ are i.i.d random variables with zero mean and a second moment. If the $l_t$ are observable as in (1), one can construct a cusum test based on

$$
\begin{aligned}
V_n(s) &:= n^{-1/2}(El_t^2)^{-1/2}\left(\sum_{t=1}^{[ns]} l_t - \left(\frac{[ns]}{n}\right)\sum_{t=1}^{n} l_t\right) \\
&\simeq \frac{[ns]}{\sqrt{n}(El_t^2)^{1/2}}(\hat{\theta}_{[ns]} - \hat{\theta}_n), \quad 0 \le s \le 1.
\end{aligned} \tag{2}
$$

However, generally the $l_t$ are unobservable, and there must be a justification for having the argument in (2). In time series models, $\{l_t\}$ usually forms a sequence of stationary martingale differences (cf. sections 3 and 4).

Now, let us consider the stationary time series $\{x_t; t = 0, \pm 1, \pm 2, \ldots\}$, and let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_J)'$ be the parameter vector, which will be examined for constancy, e.g. the mean, variance, autocovariances, etc. Here, we wish to test the following hypotheses based on the estimators $\hat{\boldsymbol{\theta}}_n$

$H_0 : \boldsymbol{\theta}$ does not change for $x_1, \ldots, x_n$ versus $H_1 :$ not $H_0$.

Let $\hat{\boldsymbol{\theta}}_k$ be the estimator of $\boldsymbol{\theta}$ based on $x_1, \ldots, x_k$. As we saw in (2), we investigate the differences $\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n$, $k = 1, \ldots, n$, for constructing a cusum test. The details are addressed below.

Suppose that $\hat{\boldsymbol{\theta}}_k$ obtained from $x_1, \ldots, x_k$, satisfies the following

$$\sqrt{k}(\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}) = \frac{1}{\sqrt{k}} \sum_{t=1}^{k} \mathbf{l}_t + \boldsymbol{\Delta}_k, \tag{3}$$

where $\mathbf{l}_t := \mathbf{l}_t(\boldsymbol{\theta}) = (l_{1,t},\ldots, l_{J,t})'$ forms stationary martingale differences with respect to a filtration $\{\mathcal{F}_t\}$, namely, for every $t$,

$$E(\mathbf{l}_t|\mathcal{F}_{t-1}) = \mathbf{0} \quad \text{a.s.}, \tag{4}$$

and $\boldsymbol{\Delta}_k = (\Delta_{1,\,k},\ldots, \Delta_{J,k})'$.

Let $\Gamma = \mathrm{Var}(\mathbf{l}_t)$ be the covariance matrix of $\mathbf{l}_t$. Assuming that $\Gamma$ is non-singular, we define the normalized martingale differences $\boldsymbol{\xi}_t := \Gamma^{-1/2}\mathbf{l}_t$. Note that $\boldsymbol{\xi}_t = (\xi_{1,t},\ldots, \xi_{J,t})'$ has uncorrelated components and satisfies (4). Thus if we put

$$\boldsymbol{\xi}_{n,t} = (\xi_{1,n,t},\ldots, \xi_{J,n,t})' := n^{-1/2}\boldsymbol{\xi}_t, \tag{5}$$

it holds that

$$\sum_{t=1}^{[ns]} \boldsymbol{\xi}_{n,t} \overset{w}{\to} \mathbf{W}_J(s) \tag{6}$$

in the $D^J[0, 1]$ space (cf. Billingsley, 1968), where $\mathbf{W}_J(s) = (W_1(s),\ldots, W_J(s))'$ denotes a $J$-dimensional standard Brownian motion, as the following conditions are satisfied (cf. Gaenssler & Haeusler, 1986, p. 311):

(1) For $j = 1, \ldots, J$ and $s \in [0, 1]$, $\displaystyle\sum_{t=1}^{[ns]} E(\xi_{j,n,t}^2|\mathcal{F}_{t-1}) \overset{P}{\to} s.$          (7)

(2) For $j = 1, \ldots, J$ and $\epsilon > 0$, $\displaystyle\sum_{t=1}^{n} E(\xi_{j,n,t}^2 I(|\xi_{j,n,t}| > \epsilon)|\mathcal{F}_{t-1}) \overset{P}{\to} 0.$       (8)

Now, suppose that for each $j = 1,\ldots, J$,

$$\max_{1 \le k \le n} \frac{\sqrt{k}}{\sqrt{n}}\, |\Delta_{j,k}| = o_P(1). \tag{9}$$

Then from (3), (6) and (9), we have that

$$\frac{[ns]}{\sqrt{n}}\, \Gamma^{-1/2}(\hat{\boldsymbol{\theta}}_{[ns]} - \boldsymbol{\theta}) = \sum_{t=1}^{[ns]} \boldsymbol{\xi}_{n,t} + \Gamma^{-1/2}\, \frac{\sqrt{[ns]}}{\sqrt{n}} \boldsymbol{\Delta}_{[ns]} \overset{w}{\to} \mathbf{W}_J(s), \tag{10}$$

and consequently,

$$\frac{[ns]}{\sqrt{n}}\, \Gamma^{-1/2}(\hat{\boldsymbol{\theta}}_{[ns]} - \hat{\boldsymbol{\theta}}_n) \overset{w}{\to} \mathbf{W}_J^{\circ}(s), \tag{11}$$

where $\mathbf{W}_J^{\circ}(s) = (W_1^{\circ}(s), \ldots, W_J^{\circ}(s))\prime$, is a $J$-dimensional standard Brownian bridge. The following is a direct result of (3)–(11).

**Theorem 1**
*Define the test statistic $T_n$ by*

$$T_n = \max_{J \le k \le n} \frac{k^2}{n} (\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n)' \Gamma^{-1} (\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n).$$

*Suppose that conditions (6) and (9) hold. Then, under $H_0$,*

$$T_n \xrightarrow{d} \sup_{0 \leq s \leq 1} \sum_{j=1}^{J} \left( W_j^{\circ}(s) \right)^2. \tag{12}$$

We reject $H_0$ if $T_n$ is large.

Using the result in (12), one can determine the critical region $(T_n \geq C_\alpha)$, given a nominal level $\alpha$, where $C_\alpha$ is the $(1 - \alpha)$-quantile point of $\sup_{0 \leq s \leq 1} \sum_{j=1}^{J} \left( W_j^{\circ}(s) \right)^2$. However, as it is not easy to calculate the critical values analytically, we provide the tables through a Monte Carlo simulation. For this task, we generate the random numbers $\epsilon_t$ following the standard normal distribution and compute the empirical quantiles based on the random variables

$$\mathcal{U}_{n,J} = \max_{1 \leq k \leq n} \sum_{j=1}^{J} \left\{ n^{-1/2} \sum_{i=1}^{k} \epsilon_{ij} - n^{-1/2} \left( \frac{k}{n} \right) \sum_{i=1}^{n} \epsilon_{ij} \right\}^2.$$

Table 1 below shows the significance levels for $\alpha = 0.01, 0.05, 0.1$ and $J = 1,\ldots, 10$, which are obtained by computing the empirical quantiles using 10,000 simulated $\mathcal{U}_{1000,M}$s.

Theorem 1 shows that the change point test in time series models can be accomplished based on any estimators provided they satisfy regularity conditions. We can say that the cusum test is widely applicable in a broad class of time series models as it constitutes the most natural non-parametric test, and some well-known estimators, such as the method of moment estimator and the Gaussian MLE, could be employed to perform a test.

## 3. Test for RCA(1) model

In this section, we focus on the RCA model. RCA models are widely used in many areas such as biology, engineering, finance and economics, and have been studied to investigate the effects of random perturbations of a dynamical system (cf. Tong, 1990). Many important properties of RCA models are reported in Nicholls & Quinn (1982) and Feigin & Tweedie (1985), some of which will be used in appropriate places.

Let $\{x_t; t = 0, \pm 1, \pm 2, \ldots\}$ be the time series of the RCA(1) model

$$x_t = (\phi + b_t)x_{t-1} + \epsilon_t, \tag{13}$$

where

$$\begin{pmatrix} b_t \\ \epsilon_t \end{pmatrix} \sim \text{i.i.d} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \omega^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \right).$$

Nicholls & Quinn (1982) showed that a sufficient condition for the strict stationarity and ergodicity of $\{x_t\}$ in (13) is $\phi^2 + \omega^2 < 1$. Here, we assume that $E\epsilon_t^{2k} < \infty$ and $E(\phi + b_t)^{2k} < 1$ for some $k$ that will be specified later, which immediately yields $Ex_t^{2k} < \infty$ (cf. Feigin & Tweedie, 1985).

Now, we consider the problem of testing for a change of the parameter vector $\boldsymbol{\theta} = (\phi, \omega^2, \sigma^2)'$ based on a (conditional) LSE $\hat{\boldsymbol{\theta}}$. Suppose that $x_1,\ldots, x_n$ are a sample from (13) and assume $x_0 = 0$. We intend to test the following hypotheses

Table 1. *Empirical $(1-\alpha)$-quantiles of $\mathcal{U}_{1000,J}$ for $J = 1,\ldots, 10$*

| | J | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 0.01 | 2.558 | 3.269 | 3.904 | 4.478 | 4.946 | 5.471 | 5.947 | 6.349 | 6.903 | 7.071 |
| 0.05 | 1.820 | 2.408 | 3.004 | 3.452 | 3.899 | 4.375 | 4.772 | 5.179 | 5.632 | 5.884 |
| 0.10 | 1.488 | 2.054 | 2.576 | 3.018 | 3.432 | 3.845 | 4.244 | 4.627 | 5.024 | 5.350 |

$H_0 : \boldsymbol{\theta} = (\phi, \omega^2, \sigma^2)'$ is constant over $x_1,\ldots, x_n$ versus

$H_1 :$ not $H_0$.

In order to construct a cusum test, consider the estimators $\hat{\boldsymbol{\theta}}_k = (\hat{\phi}_k, \hat{\omega}_k^2, \hat{\sigma}_k^2)'$ based on $x_1,\ldots, x_k$, $k = 1,\ldots, n$. The estimator $\hat{\phi}_k$ of $\phi$ is defined as the minimizer of $\sum_{t=1}^{k}(x_t - \phi x_{t-1})^2$, and the estimators $\hat{\omega}_k^2$ and $\hat{\sigma}_k^2$ are defined as the minimizers of $\sum_{t=1}^{k}(\hat{u}_{k,t}^2 - \omega^2 x_{t-1}^2 - \sigma^2)^2$, where $\hat{u}_{k,t} = x_t - \hat{\phi}_k x_{t-1}$, by noticing the equation

$$E(u_t^2|\mathcal{F}_{t-1}) = \omega^2 x_{t-1}^2 + \sigma^2 \quad \text{under } H_0,$$

where $u_t = x_t - \phi x_{t-1}$ and $\mathcal{F}_t = \sigma(\epsilon_s, b_s; s \le t)$. Then, $\hat{\boldsymbol{\theta}}_k$ is written as

$$\hat{\boldsymbol{\theta}}_k = \begin{pmatrix} \hat{\phi}_k \\ \hat{\omega}_k^2 \\ \hat{\sigma}_k^2 \end{pmatrix} = \begin{pmatrix} \dfrac{\sum_{t=1}^{k} x_{t-1}x_t}{\sum_{t=1}^{k} x_{t-1}^2} \\[2ex] \dfrac{\sum_{t=1}^{k}(x_{t-1}^2 - m_{2,k})\hat{u}_{k,t}^2}{\sum_{t=1}^{k}(x_{t-1}^2 - m_{2,k})^2} \\[2ex] k^{-1}\sum_{t=1}^{k}\hat{u}_{k,t}^2 - \hat{\omega}_k^2 m_{2,k} \end{pmatrix},$$

where $m_{2,k} = k^{-1}\sum_{t=1}^{k} x_{t-1}^2$.

Suppose that the null hypothesis $H_0$ is true. In order to apply the procedure in section 2, we decompose $\hat{\boldsymbol{\theta}}_k$ into the sum of martingale differences with respect to $\{\mathcal{F}_t\}$ and negligible terms as in (3)

$$\sqrt{k}\begin{pmatrix} \hat{\phi}_k - \phi \\ \hat{\omega}_k^2 - \omega^2 \\ \hat{\sigma}_k^2 - \sigma^2 \end{pmatrix} = \frac{1}{\sqrt{k}}\begin{pmatrix} \sum_{t=1}^{k} l_{1,t} \\ \sum_{t=1}^{k} l_{2,t} \\ \sum_{t=1}^{k} l_{3,t} \end{pmatrix} + \begin{pmatrix} \Delta_{1,k} \\ \Delta_{2,k} \\ \Delta_{3,k} \end{pmatrix}, \tag{14}$$

where

$$l_{1,t} = \frac{x_{t-1}u_t}{Ex_1^2},$$

$$l_{2,t} = \frac{(x_{t-1}^2 - Ex_1^2)(u_t^2 - \omega^2 x_{t-1}^2 - \sigma^2)}{E(x_1^2 - Ex_1^2)^2},$$

$$l_{3,t} = u_t^2 - \omega^2 x_{t-1}^2 - \sigma^2 - l_{2,t}Ex_1^2,$$

$$\Delta_{1,k} = \frac{Ex_1^2 - m_{2,k}}{Ex_1^2 m_{2,k}}\frac{1}{\sqrt{k}}\sum_{t=1}^{k} x_{t-1}u_t,$$

$$\Delta_{2,k} = \frac{1}{\sqrt{k}}\sum_{t=1}^{k}(u_t^2 - \omega^2 x_{t-1}^2 - \sigma^2)\left(\frac{x_{t-1}^2 - m_{2,k}}{k^{-1}\sum_{t=1}^{k}(x_{t-1}^2 - m_{2,k})^2} - \frac{x_{t-1}^2 - Ex_1^2}{E(x_1^2 - Ex_1^2)^2}\right)$$
$$+ \sqrt{k}\sum_{t=1}^{k}\frac{(x_{t-1}^2 - m_{2,k})}{\sum_{t=1}^{k}(x_{t-1}^2 - m_{2,k})^2}\left(2(\phi - \hat{\phi}_k)x_{t-1}u_t + (\phi - \hat{\phi}_k)^2 x_{t-1}^2\right),$$

$$\Delta_{3,k} = -\Delta_{2,k}Ex_1^2 + \frac{1}{\sqrt{k}}\sum_{t=1}^{k}\left(2(\phi - \hat{\phi}_k)x_{t-1}u_t + (\phi - \hat{\phi}_k)^2 x_{t-1}^2\right)$$
$$- \frac{1}{\sqrt{k}}\sum_{t=1}^{k}(x_{t-1}^2 - Ex_1^2)(\hat{\omega}_k^2 - \omega^2).$$

It is obvious that for each $j = 1, 2, 3$, $\{l_{j,t}\}$ in (14) forms a sequence of martingale differences with respect to $\{\mathcal{F}_t\}$. Furthermore, from the stationarity and ergodicity of $\{x_t\}$, $\{l_{j,t}\}$ is also stationary and ergodic.

Putting $\Gamma_{ij} = E(l_{i,t}\, l_{j,t})$ and $\Gamma = \left(\Gamma_{ij}\right)_{i,j=1}^{3}$, we can see that the random variables $\xi_{j,n,t}$, $j = 1, 2, 3$, as defined in (5), are strictly stationary and ergodic martingale differences. Thus, $\{\xi_{j,n,t}\}$ satisfies conditions (7) and (8), and hence,

$$\left(\sum_{t=1}^{[ns]}\xi_{1,n,t}, \sum_{t=1}^{[ns]}\xi_{2,n,t}, \sum_{t=1}^{[ns]}\xi_{3,n,t}\right)' \xrightarrow{w} \mathbf{W}_3(s). \tag{15}$$

In view of (15) and (9), it suffices to show that

$$\max_{1 \le k \le n} \frac{\sqrt{k}}{\sqrt{n}}\, |\Delta_{j,k}| = o_P(1), \quad j = 1, 2, 3, \tag{16}$$

to obtain the convergence result in theorem 2. The following is the main result of this section.

**Theorem 2**
*Suppose that $E\epsilon_1^{16} < \infty$ and $E(\phi + b_1)^{16} < 1$. Then under $H_0$, as $n \to \infty$,*

$$\frac{[ns]}{\sqrt{n}}\, \Gamma^{-1/2}(\hat{\boldsymbol{\theta}}_{[ns]} - \hat{\boldsymbol{\theta}}_n) \xrightarrow{w} \mathbf{W}_3^{\circ}(s).$$

In view of the result of theorem 2, we can construct the test statistic

$$T_n = \max_{1 \le k \le n} \frac{k^2}{n}\, (\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n)'\hat{\Gamma}^{-1}(\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n),$$

where $\hat{\Gamma}$ is a consistent estimator of $\Gamma$. We reject $H_0$ at the level $\alpha$ if $T_n \ge C_\alpha$, where $C_\alpha$ is the $(1 - \alpha)$-quantile point of $\sup_{0 \le s \le 1} \sum_{j=1}^{3}\left(W_j^{\circ}(s)\right)^2$.

In fact, one can calculate

$$\Gamma_{11} = \frac{\omega^2 Ex_1^4 + \sigma^2 Ex_1^2}{\left(Ex_1^2\right)^2},$$

$$\Gamma_{22} = \left(Ex_1^4 - (Ex_1^2)^2\right)^{-2}\Big((Eb_1^4 - \omega^4)(Ex_1^8 - 2Ex_1^2Ex_1^6 + (Ex_1^2)^2Ex_1^4)$$
$$+ 4\omega^2\sigma^2(Ex_1^6 - 2Ex_1^2Ex_1^4 + (Ex_1^2)^3) + (E\epsilon_1^4 - \sigma^4)(Ex_1^4 - (Ex_1^2)^2)\Big),$$

$$\Gamma_{33} = \left(Eb_1^4 - \omega^4\right)\left(Ex_1^4 - \frac{2Ex_1^2(Ex_1^6 - Ex_1^2Ex_1^4)}{Ex_1^4 - (Ex_1^2)^2}\right)$$
$$- 4\omega^2\sigma^2 Ex_1^2 + E\epsilon_1^4 - \sigma^4 + (Ex_1^2)^2\Gamma_{22},$$

$$\Gamma_{12} = \frac{Eb_1^3 Ex_1^6 - Eb_1^3 Ex_1^2 Ex_1^4 + E\epsilon_1^3 Ex_1^3}{Ex_1^2 Ex_1^4 - \left(Ex_1^2\right)^3},$$

$$\Gamma_{13} = \frac{-Eb_1^3 Ex_1^2 Ex_1^6 + Eb_1^3\left(Ex_1^4\right)^2 - E\epsilon_1^3 Ex_1^2 Ex_1^3}{Ex_1^2 Ex_1^4 - \left(Ex_1^2\right)^3},$$

$$\Gamma_{23} = \frac{(Eb_1^4 - \omega^4)(Ex_1^6 - Ex_1^2 Ex_1^4)}{Ex_1^4 - \left(Ex_1^2\right)^2} + 4\omega^2\sigma^2 - Ex_1^2\Gamma_{22}.$$

Therefore, in order to obtain $\hat{\Gamma}$, one should estimate $E\epsilon_t^3$, $Eb_t^3$, $E\epsilon_t^4$ and $Eb_t^4$. For the estimators of $E\epsilon_t^3$ and $Eb_t^3$, we employ the LSEs to minimize $\sum_{t=1}^{n}(\hat{u}_t^3 - x_{t-1}^3 Eb_t^3 + E\epsilon_t^3)^2$ in view of the equation $E(u_t^3|\mathcal{F}_{t-1}) = x_{t-1}^3 Eb_t^3 + E\epsilon_t^3$. The estimators for $E\epsilon_t^4$ and $Eb_t^4$ are similarly obtained. Plugging those estimators and $n^{-1}\sum_{t=1}^{n} x_t^k$, $k = 2, 3, 4, 6, 8$, into $\Gamma_{ij}$, one can get a consistent estimator $\hat{\Gamma}$ of $\Gamma$.

Now, we prove (16). The following lemma is needed for later work. We state it without proof.

**Lemma 1**

*If a double array of random variables $\{x_{n,k}: 1 \le k \le n, n \ge 1\}$ satisfies the conditions*

(i) *For each $N \ge 1$, $\max_{1 \le k \le N}|x_{n,k}| = o_P(1)$ as $n \to \infty$.*

(ii) *For any $\epsilon > 0$ and $\delta > 0$, there exist $N$ and $L$, such that $P(\max_{N \le k \le n}|x_{n,k}| > \epsilon) < \delta$ for all $n \ge L$,*

*then $\max_{1 \le k \le n}|x_{n,k}| = o_P(1)$.*

The following lemma is concerned with the negligibility of $\Delta_{1,k}$.

**Lemma 2**

*Under the conditions in theorem 2,*

$$\max_{1 \le k \le n}\left|\frac{\sqrt{k}}{\sqrt{n}}\Delta_{1,k}\right| = o_P(1).$$

*Proof.* First, note that

$$\frac{1}{\sqrt{n}}\max_{1 \le k \le n}\left|\sqrt{k}\Delta_{1,k}\right| \le \max_{1 \le k \le n}\left|\frac{\sum_{t=1}^{k} x_{t-1}u_t}{\sqrt{2k \log\log k}}\right|\frac{1}{\sqrt{n}}\max_{1 \le k \le n}\left|\frac{\sqrt{2k \log\log k}(m_{2,k} - Ex_1^2)}{Ex_1^2 m_{2,k}}\right|.$$

As $\{x_{t-1}u_t\}$ are stationary and ergodic, we can see that

$$\max_{1 \le k \le n}\left|\frac{\sum_{t=1}^{k} x_{t-1}u_t}{\sqrt{2k \log\log k}}\right| = O_P(1),$$

by the law of the iterated logarithm (LIL) for martingales (cf. Stout, 1970). Note that for each $N \ge 1$, as $n \to \infty$,

$$\frac{1}{\sqrt{n}}\max_{1 \le k \le N}\left|\frac{\sqrt{2k \log\log k}(m_{2,k} - Ex_1^2)}{Ex_1^2 m_{2,k}}\right| = o_P(1).$$

Hence, in order to obtain

$$\frac{1}{\sqrt{n}}\max_{1 \le k \le n}\left|\frac{\sqrt{2k \log\log k}(m_{2,k} - Ex_1^2)}{Ex_1^2 m_{2,k}}\right| = o_P(1),$$

it suffices to show that condition (ii) of lemma 1 holds.

For any $\epsilon > 0$, $\delta > 0$ and $N \ge 1$, observe that

$$R_{n,N} := P\left(\frac{1}{\sqrt{n}}\max_{N \le k \le n}\left|\frac{\sqrt{2k \log\log k}(m_{2,k} - Ex_1^2)}{Ex_1^2 m_{2,k}}\right| > \epsilon\right)$$

$$\le P\left(\frac{1}{\sqrt{n}}\max_{N \le k \le n} 2 \log\log k \max_{N \le k \le n}\left|\frac{2\sum_{t=1}^{k}(x_{t-1}^2 - Ex_1^2)}{\sqrt{2k \log\log k}(Ex_1^2)^2}\right| > \epsilon\right) +$$

$$\sum_{k=N}^{\infty} P\left(k^{-1}\left|\sum_{t=1}^{k}(x_{t-1}^2 - Ex_1^2)\right| \ge \frac{Ex_1^2}{2}\right)$$

$$=: I_n + II_N.$$

Note that by the LIL,

$$\max_{N \le k \le n} \left| \frac{\sum_{t=1}^{k}(x_{t-1}^2 - Ex_1^2)}{\sqrt{2k \log \log k}} \right| = O_P(1),$$

so that $I_n < \delta/2$ for sufficiently large $n$, say $n \ge L$.

Meanwhile, by Markov's inequality, we have

$$\sum_{k=N}^{\infty} P\left( k^{-1} \left| \sum_{t=1}^{k}(x_{t-1}^2 - Ex_1^2) \right| \ge \frac{Ex_1^2}{2} \right) \le \left( \frac{Ex_1^2}{2} \right)^{-4} \sum_{k=N}^{\infty} E\left( k^{-1} \sum_{t=1}^{k}(x_{t-1}^2 - Ex_1^2) \right)^4.$$

Therefore, we can choose $N$ such that $II_N < \delta/2$, as

$$\sum_{k=1}^{\infty} E\left( k^{-1} \sum_{t=1}^{k}(x_{t-1}^2 - Ex_1^2) \right)^4 < \infty.$$

This yields that $R_{n,N} < \delta$ for all $n \ge L$, and the lemma is established.

**Lemma 3**
*Under the conditions in theorem 2,*

$$\max_{1 \le k \le n} \left| \frac{\sqrt{k}}{\sqrt{n}} \Delta_{2,k} \right| = o_P(1).$$

*Proof.* Write

$$\frac{\sqrt{k}}{\sqrt{n}} \Delta_{2,k} = \frac{1}{\sqrt{n}} \sum_{t=1}^{k}(u_t^2 - \omega^2 x_{t-1}^2 - \sigma^2) \left( \frac{x_{t-1}^2 - m_{2,k}}{k^{-1}\sum_{t=1}^{k}(x_{t-1}^2 - m_{2,k})^2} - \frac{x_{t-1}^2 - Ex_1^2}{E(x_1^2 - Ex_1^2)^2} \right)$$
$$+ \frac{k}{\sqrt{n}} \sum_{t=1}^{k} \frac{(x_{t-1}^2 - m_{2,k})}{\sum_{t=1}^{k}(x_{t-1}^2 - m_{2,k})^2} \left( 2(\phi - \hat{\phi}_k)x_{t-1}u_t + (\phi - \hat{\phi}_k)^2 x_{t-1}^2 \right)$$
$$= I_k + II_k.$$

For $I_k$, observe that

$$\max_{1 \le k \le n} |I_k| \le \max_{1 \le k \le n} \left| \sum_{t=1}^{k} \frac{(u_t^2 - \omega^2 x_{t-1}^2 - \sigma^2)(x_{t-1}^2 - Ex_1^2)}{\sqrt{2k \log \log k}} \right|$$
$$\times \frac{1}{\sqrt{n}} \max_{1 \le k \le n} \left| \frac{\sqrt{2k \log \log k}\left( E(x_1^2 - Ex_1^2)^2 - k^{-1}\sum_{t=1}^{k}(x_{t-1}^2 - m_{2,k})^2 \right)}{E(x_1^2 - Ex_1^2)^2 k^{-1}\sum_{t=1}^{k}(x_{t-1}^2 - m_{2,k})^2} \right|$$
$$+ \max_{1 \le k \le n} \left| \sum_{t=1}^{k} \frac{(u_t^2 - \omega^2 x_{t-1}^2 - \sigma^2)}{\sqrt{2k \log \log k}} \right| \times \frac{1}{\sqrt{n}} \max_{1 \le k \le n} \left| \frac{\sqrt{2k \log \log k}(Ex_1^2 - m_{2,k})}{k^{-1}\sum_{t=1}^{k}(x_{t-1}^2 - m_{2,k})^2} \right|$$

As $\{(u_t^2 - \omega^2 x_{t-1}^2 - \sigma^2)(x_{t-1}^2 - Ex_1^2)\}$ and $\{u_t^2 - \omega^2 x_{t-1}^2 - \sigma^2\}$ are stationary and ergodic, similar to the arguments in the proof of lemma 2, we can see that the first and third terms are $O_P$ (1) and the second and fourth terms are $o_P$ (1). Therefore, we can conclude that $\max_{1 \le k \le n}|I_k| = o_P$ (1).

Now, we prove $\max_{1 \leq k \leq n} |II_k| = o_P(1)$. Observe that

$$
\begin{aligned}
\max_{1 \leq k \leq n} |II_k| \leq {} & 2 \max_{1 \leq k \leq n} \left| \frac{k}{\sqrt{n}} \sum_{t=1}^{k} \left( \frac{(x_{t-1}^2 - m_{2,k}) x_{t-1} u_t}{\sum_{t=1}^{k} (x_{t-1}^2 - m_{2,k})^2} \right) (\phi - \hat{\phi}_k) \right| \\
& + \max_{1 \leq k \leq n} \left| \frac{k}{\sqrt{n}} \sum_{t=1}^{k} \left( \frac{(x_{t-1}^2 - m_{2,k}) x_{t-1}^2}{\sum_{t=1}^{k} (x_{t-1}^2 - m_{2,k})^2} \right) (\phi - \hat{\phi}_k)^2 \right| \\
= {} & II_{k1} + II_{k2}.
\end{aligned}
$$

In order to show $II_{k1} = o_P(1)$, it suffices to show that

$$
\frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^{k} (x_{t-1}^2 - m_{2,k}) x_{t-1} u_t \right| |\phi - \hat{\phi}_k| = o_P(1). \tag{17}
$$

First, note that by the LIL,

$$
\max_{1 \leq k \leq n} \left| \frac{\sum_{t=1}^{k} (x_{t-1}^2 - m_{2,k}) x_{t-1} u_t}{\sqrt{2k \log \log k}} \right| = O_P(1). \tag{18}
$$

Furthermore, from the relationship $|\hat{\phi}_k - \phi| \leq |\sum_{t=1}^{k} x_{t-1} u_t| / (k E x_1^2) + |\Delta_{1,k}| / \sqrt{k}$,

$$
\frac{1}{\sqrt{n}} \max_{N \leq k \leq n} \sqrt{2k \log \log k} |\hat{\phi}_k - \phi| = o_P(1).
$$

This together with (18) implies (17). Meanwhile, noticing

$$
\frac{1}{\sqrt{n}} \frac{\sum_{t=1}^{k} (x_{t-1}^2 - m_{2,k}) m_{2,k}}{k^{-1} \sum_{t=1}^{k} (x_{t-1}^2 - m_{2,k})^2} (\phi - \hat{\phi}_k)^2 = 0,
$$

we can see that the term $II_{k2}$ reduces to $\max_{1 \leq k \leq n} \left| \frac{k}{\sqrt{n}} (\hat{\phi}_k - \phi)^2 \right|$. As for any $\epsilon > 0$,

$$
P\left( \max_{1 \leq k \leq n} \left| \frac{k}{\sqrt{n}} (\hat{\phi}_k - \phi)^2 \right| > \epsilon \right) = P\left( \max_{1 \leq k \leq n} \frac{2 \log \log k}{\sqrt{n}} \left| \frac{\sum_{t=1}^{k} x_{t-1} u_t / \sqrt{2k \log \log k}}{k^{-1} \sum_{t=1}^{k} x_{t-1}^2} \right|^2 > \epsilon \right) = o(1),
$$

we have $II_{k2} = o_P(1)$. This completes the proof.

It remains to show the asymptotic negligibility of $\Delta_{3,k}$ to complete the proof of theorem 2.

### Lemma 4
*Under the conditions in theorem* 2,

$$
\max_{1 \leq k \leq n} \left| \frac{\sqrt{k}}{\sqrt{n}} \Delta_{3,k} \right| = o_P(1).
$$

*Proof.* Split $\max_{1 \leq k \leq n} \frac{\sqrt{k}}{\sqrt{n}} \Delta_{3,k}$ into the following four terms

$$
\begin{aligned}
\max_{1 \leq k \leq n} \left| \frac{\sqrt{k}}{\sqrt{n}} \Delta_{3,k} \right| \leq {} & \max_{1 \leq k \leq n} \left| \frac{\sqrt{k}}{\sqrt{n}} \Delta_{2,k} \right| E x_1^2 \\
& + \max_{1 \leq k \leq n} \left| 2(\phi - \hat{\phi}_k) \frac{1}{\sqrt{n}} \sum_{t=1}^{k} x_{t-1} u_t \right| + \max_{1 \leq k \leq n} \left| (\phi - \hat{\phi}_k)^2 \frac{1}{\sqrt{n}} \sum_{t=1}^{k} x_{t-1}^2 \right| \\
& + \max_{1 \leq k \leq n} \left| \frac{\sum_{t=1}^{k} (x_{t-1}^2 - E x_1^2)}{\sqrt{2k \log \log k}} \right| \max_{1 \leq k \leq n} \left| \frac{\sqrt{2k \log \log k}}{\sqrt{n}} (\hat{\omega}_k^2 - \omega^2) \right|.
\end{aligned}
$$

Similar to the proof of lemma 3, we can see that each term is $o_P(1)$. Following some algebra, the details of which we omit, we establish the lemma.

## 4. Test for autocovariance function

In this section, we consider the problem of testing for a change of autocovariance function in infinite order moving average processes in light of the results of section 2. Let $\{x_t; t = 0, \pm 1, \pm 2, \ldots\}$ be a stationary linear process of the form

$$x_t = \sum_{i=0}^{\infty} a_i \epsilon_{t-i}, \tag{19}$$

where the real sequence $\{a_i\}$ satisfies the summability condition $\sum_{i=0}^{\infty} i|a_i| < \infty$ and $\epsilon_t$ are i.i.d random variables with mean 0, variance $\sigma_\epsilon^2$, and $E|\epsilon_1|^{4\lambda} < \infty$ for some $\lambda > 1$. Assume that $x_1, \ldots, x_n$ are observed, and denote the autocovariance at lag $h$ by $\gamma(h)$. As an estimate of $\gamma(h)$, we use

$$\hat{\gamma}_n(h) = \frac{1}{n} \sum_{t=1}^{n-h} (x_t - \bar{x}_n)(x_{t+h} - \bar{x}_n), \quad \bar{x}_n = \frac{1}{n} \sum_{t=1}^{n} x_t.$$

The following result is to show that the weak convergence result in (11) holds.

**Theorem 3**

*Let*

$$\Lambda_n(s) = \left( \frac{[ns]}{\sqrt{n}} \left( \hat{\gamma}_{[ns]}(0) - \hat{\gamma}_n(0) \right), \ldots, \frac{[ns]}{\sqrt{n}} \left( \hat{\gamma}_{[ns]}(m) - \hat{\gamma}_n(m) \right) \right)', \quad 0 \leq s \leq 1.$$

*Then under $H_0$, where no changes are assumed to occur in the autocovariance function, we have*

$$\Lambda_n(s)' \Gamma^{-1} \Lambda_n(s) \xrightarrow{w} \sum_{j=0}^{m} \left( W_j^\circ(s) \right)^2,$$

*where $\Gamma$ is the $(m+1) \times (m+1)$ matrix whose $(i,j)$th entry is*

$$\Gamma_{ij} = \kappa_4 \gamma(i)\gamma(j) + \sum_{r=-\infty}^{\infty} (\gamma(i+r)\gamma(j+r) + \gamma(i-r)\gamma(j+r)), \quad i, j = 0, \ldots, m,$$

*and $\kappa_4$ is the kurtosis of $\epsilon_1$.*

As $\Gamma$ is unknown, we should replace it by a consistent estimator $\hat{\Gamma}$. Now we assume that $\{x_t\}$ in (19) can be rewritten as

$$x_t = \sum_{j=1}^{\infty} \pi_j x_{t-j} + \epsilon_t, \tag{20}$$

where $\pi(z) := 1 - \sum_{j=1}^{\infty} \pi_j z^j$ is analytic on an open set containing the unit disk in the complex plane, and have no zeros in the unit disk. Notice that $\{x_t\}$ can be rewritten as in (19), and covers stationary and invertible ARMA processes. We introduce a sequence of positive integers $\{h_n\}$, such that as $n \to \infty$,

$$h_n \to \infty \text{ and } h_n = O(n^\beta) \text{ for some } \beta \in \left( 0, \frac{(\lambda - 1)}{2\lambda} \right).$$

Then if $\hat{\kappa}_4$ is a consistent estimator of $\kappa_4$, we have

$$\hat{\Gamma}_{ij} \xrightarrow{P} \Gamma_{ij}, \tag{21}$$

where

$$\hat{\Gamma}_{ij} = \hat{\kappa}_4 \hat{\gamma}_n(i) \hat{\gamma}_n(j) + \sum_{r=-h_n}^{h_n} (\hat{\gamma}_n(i+r)\hat{\gamma}_n(j+r) + \hat{\gamma}_n(i-r)\hat{\gamma}_n(j+r)), \quad i, j = 0, \ldots, m.$$

The argument in (21) can be readily proved by using lemma 4.2 (ii) of Lee (1996, p. 2239). Note that a consistent $\hat{\kappa}_4$ can be obtained by calculating residuals $\hat{\epsilon}_t$ via fitting a long AR($q$) model to observations (cf. Lee & Wei, 1999), viz., $\hat{\kappa}_4 = n^{-1} \sum_{t=1}^{n} \hat{\epsilon}_t^4 / (n^{-1} \sum_{t=1}^{n} \hat{\epsilon}_t^2)^2 - 3$. A typical example of $q$ is ($\log n$)$^2$. From theorem 3 and (21), we obtain the following result.

**Theorem 4**
*Under* $H_0$,

$$\Lambda_n(s)' \hat{\Gamma}^{-1} \Lambda_n(s) \xrightarrow{w} \sum_{j=0}^{m} \left( W_j^{\circ}(s) \right)^2.$$

Theorem 4 ensures the Brownian bridge result for the cusum test statistic. The test statistic is defined as

$$\Lambda_n := \max_{m+1 \leq k \leq n} \Lambda_n(k/n)' \hat{\Gamma}^{-1} \Lambda_n(k/n).$$

Now, we prove theorem 3.

*Proof of Theorem 3.* Note that

$$\sqrt{k}(\hat{\gamma}_k(h) - \gamma(h)) = \frac{1}{\sqrt{k}} \sum_{t=1}^{k} (x_t x_{t+h} - \gamma(h)) + \delta_{h,k},$$

where

$$\delta_{h,k} = -\frac{1}{\sqrt{k}} \bar{x}_k \sum_{t=1}^{k} x_{t+h} - \frac{1}{\sqrt{k}} \sum_{t=k-h+1}^{k} (x_t - \bar{x}_k)(x_{t+h} - \bar{x}_k).$$

In view of Phillips & Solo (1992, p. 980), we can write that

$$\begin{aligned}
x_t x_{t+h} &= f_h(B)\epsilon_t^2 + \sum_{r=1}^{\infty} (f_{h+r}(B)\epsilon_{t-r}\epsilon_t + f_{h-r}(B)\epsilon_{t+r}\epsilon_t) \\
&= f_h(1)\epsilon_t^2 + \sum_{r=1}^{\infty} (f_{h+r}(1)\epsilon_{t-r}\epsilon_t + f_{h-r}(1)\epsilon_{t+r}\epsilon_t) \\
&\quad - (1-B)\tilde{f}_h(B)\epsilon_t^2 - (1-B) \sum_{r=1}^{\infty} (\tilde{f}_{h+r}(B)\epsilon_{t-r}\epsilon_t + \tilde{f}_{h-r}(1)\epsilon_{t+r}\epsilon_t),
\end{aligned}$$

where $B$ is a back-shift operator,

$$f_j(B) = \sum_{i=0}^{\infty} f_{ji} B^i = \sum_{i=0}^{\infty} a_i a_{i+j} B^i$$

and

$$\tilde{f}_j(B) = \sum_{i=0}^{\infty} \tilde{f}_{ji} B^i = \sum_{i=0}^{\infty} \left( \sum_{l=i+1}^{\infty} f_{jl} \right) B^i = \sum_{i=0}^{\infty} \left( \sum_{l=i+1}^{\infty} a_l a_{l+j} \right) B^i.$$

As $\gamma(h) = f_h(1)\sigma_\epsilon^2$, we can write

$$\frac{1}{\sqrt{k}} \sum_{t=1}^{k} (x_t x_{t+h} - \gamma(h)) = \frac{1}{\sqrt{k}} \sum_{t=1}^{k} l_{h,t} + \sum_{j=1}^{3} \Delta_{h,k,j},$$

where

$$l_{h,t} = f_h(1)(\epsilon_t^2 - \sigma_\epsilon^2) + \sum_{r=1}^{\infty} (f_{h+r}(1) + f_{h-r}(1))\epsilon_{t-r}\epsilon_t,$$

$$\Delta_{h,k,1} = \sum_{r=1}^{\infty} f_{h-r} \left( \frac{1}{\sqrt{k}} \sum_{t=1}^{k} (\epsilon_{t+r}\epsilon_t - \epsilon_t\epsilon_{t-r}) \right),$$

$$\Delta_{h,k,2} = -\frac{1}{\sqrt{k}} \tilde{f}_h(B)\epsilon_k^2 + \frac{1}{\sqrt{k}} \tilde{f}_h(B)\epsilon_0^2 - \frac{1}{\sqrt{k}} \sum_{r=1}^{\infty} (\tilde{f}_{h+r}(B)\epsilon_{k-r}\epsilon_k + \tilde{f}_{h-r}(1)\epsilon_{k+r}\epsilon_k)$$

$$+ \frac{1}{\sqrt{k}} \sum_{r=1}^{\infty} (\tilde{f}_{h+r}(B)\epsilon_{-r}\epsilon_0 + \tilde{f}_{h-r}(1)\epsilon_r\epsilon_0),$$

$$\Delta_{h,k,3} = \delta_{h,k}.$$

First, observe that

$$\Gamma^{-1/2} \left( \frac{1}{\sqrt{n}} \sum_{t=1}^{[ns]} l_{0,t}, \ldots, \frac{1}{\sqrt{n}} \sum_{t=1}^{[ns]} l_{m,t} \right)' \xrightarrow{w} \mathbf{W}_{m+1}(s), \tag{22}$$

which can be shown similarly to (6). Secondly, note that

$$\max_{1 \le k \le n} \frac{\sqrt{k}}{\sqrt{n}} |\Delta_{h,k,1}| = o_P(1), \tag{23}$$

as

$$\max_{1 \le k \le n} \frac{\sqrt{k}}{\sqrt{n}} |\Delta_{h,k,1}| \le \frac{1}{\sqrt{n}} \max_{1 \le k \le n} \left| \sum_{r=1}^{k-1} f_{h-r}(1) \sum_{t=k-r+1}^{k} \epsilon_{t+r}\epsilon_t \right|$$

$$+ \frac{1}{\sqrt{n}} \max_{1 \le k \le n} \left| \sum_{r=1}^{k-1} f_{h-r}(1) \sum_{t=1}^{r} \epsilon_t\epsilon_{t-r} \right|$$

$$+ \frac{1}{\sqrt{n}} \max_{1 \le k \le n} \left| \sum_{r=-k}^{\infty} f_{h-r}(1) \sum_{t=1}^{k} (\epsilon_{t+r}\epsilon_t - \epsilon_t\epsilon_{t-r}) \right|$$

$$= O_P(n^{-1/4}),$$

where we have used Minkowski's inequality and the fact that $\sum_{r=1}^{\infty} r|f_{h-r}(1)| < \infty$.

Thirdly, we have that

$$\max_{1 \le k \le n} \frac{\sqrt{k}}{\sqrt{n}} |\Delta_{h,k,2}| = o_P(1), \tag{24}$$

as

$$\max_{1 \le k \le n} \frac{1}{\sqrt{n}} |\tilde{f}_h(B)\epsilon_k^2| + \max_{1 \le k \le n} \frac{1}{\sqrt{n}} \left| \sum_{r=1}^{\infty} (\tilde{f}_{h+r}(B)\epsilon_{k-r}\epsilon_k + \tilde{f}_{h-r}(1)\epsilon_{k+r}\epsilon_k) \right| = o_P(1).$$

Finally, we have that

$$\max_{1\leq k\leq n} \frac{\sqrt{k}}{\sqrt{n}} |\Delta_{h,k,3}| = o_P(1), \tag{25}$$

as

$$\max_{1\leq k\leq n} \frac{k}{\sqrt{n}} \bar{x}_k^2 \leq \frac{2\log\log n}{\sqrt{n}} \max_{1\leq k\leq n} \left| \frac{1}{\sqrt{2k\log\log k}} \sum_{t=1}^{k} x_k \right|^2 = o_P(1),$$

which is due to theorem 4.3 of Phillips & Solo (1992, p. 977), so that

$$\max_{1\leq k\leq n} \frac{\sqrt{k}}{\sqrt{n}} |\Delta_{h,k,3}| \leq \max_{1\leq k\leq n} \frac{1}{\sqrt{n}} \left| \bar{x}_k \sum_{t=1}^{k} x_{t+h} \right| + \max_{1\leq k\leq n} \frac{1}{\sqrt{n}} \left| \sum_{t=k-h+1}^{k} (x_t - \bar{x}_k)(x_{t+h} - \bar{x}_k) \right|$$

$$\leq \max_{1\leq k\leq n} \frac{k}{\sqrt{n}} \bar{x}_k^2 + \max_{1\leq k\leq n} \frac{1}{\sqrt{n}} \left| \bar{x}_k \sum_{t=k+1}^{k+h} x_t \right| + \max_{1\leq k\leq n} \frac{1}{\sqrt{n}} \left| \bar{x}_k \sum_{t=1}^{h} x_t \right|$$

$$+ \max_{1\leq k\leq n} \frac{1}{\sqrt{n}} \left| \sum_{t=k-h+1}^{k} (x_t - \bar{x}_k)(x_{t+h} - \bar{x}_k) \right| = o_P(1).$$

Combining (22)–(25), we establish the theorem.

## 5. Simulation results

In this section, we evaluate the performance of the test statistics $T_n$ and $\Lambda_n$ through a simulation study. In particular, $T_n$ and $\Lambda_n$ are compared with Nyblom's (1989) test statistic $T_n'$, which equals $\hat{L}/n^2$ in Nyblom, and Picard's (1985, p. 843) test statistic $Z_n$, respectively. For $T_n'$, we employed the MLE for the parameters in the RCA(1) model, which can be found in Quinn & Nicholls (1981). The empirical sizes and powers are calculated at the nominal level 0.1 in both cases. In each simulation, 100 initial observations are discarded to remove initialization effects. First, we deal with $T_n$ in section 3. For the empirical sizes of $T_n$, sets of 200, 400, 600 and 800 observations are generated from the AR(1) model with $\phi = 0$, 0.3 and 0.5, and also from the RCA(1) model with $\phi = 0$, 0.3, where $\{b_t\}$ are normal random variables with mean zero and $\omega^2 = 0.1$. In both cases, $\{\epsilon_t\}$ is generated from normal random variables with zero mean and $\sigma^2 = 1$. In order to see the power, we consider the alternatives as follows:

(i) Change from the AR(1) model to the RCA(1) model with $\omega^2 = 0.5$, and
(ii) Change from the RCA model with $\omega^2 = 0.1$ to the one with $\omega^2 = 0.5$.

In each case, we consider two cases: (i) $\sigma^2$ remains equal to 1, and (ii) $\sigma^2$ changes from 1 to 2. Further, for each simulation we take into consideration the cases where $\phi$ changes from 0 to 0, 0.3 and 0.5, from 0.3 to 0.3 and 0.5, and from 0.5 to 0.5, respectively. As in Table 3, we use the symbol $a \rightarrow b$ to denote the change from $a$ to $b$. All these changes are assumed to occur at the centre of the observations. The critical value at $\alpha = 0.1$ is 2.576, which is obtained from Table 1.

The figures in Tables 2–6 indicate the proportion of the number of rejections of the null hypothesis, '$H_0$: No changes occur in $\theta = (\phi, \omega^2, \sigma^2)'$', out of 500 repetitions: the figures within parentheses are for $T_n'$. From Table 2, we can see that neither $T_n$ nor $T_n'$ produces severe size distortions. In the meantime, from Tables 3–6, we can see that the powers of $T_n$ and $T_n'$ are fairly good in all cases. In particular, there is a tendency for $T_n$ to produce better powers than $T_n'$ to some degree. As expected, the powers increase remarkably as more parameters experience changes. For example, as seen in Tables 4 and 6, the powers are close to 1 when more than two parameters change.

Table 2. *Empirical sizes of $T_n$ and $T_n'$ at nominal level 0.1 when $\sigma^2 = 1$. The figures for $T_n'$ are within parentheses*

| $\omega^2$ | | 0 | | | 0.1 | |
|---|---|---|---|---|---|---|
| $\phi$ | | 0.0 | 0.3 | 0.5 | 0.0 | 0.3 |
| | 200 | 0.118 (0.058) | 0.160 (0.052) | 0.128 (0.052) | 0.108 (0.074) | 0.150 (0.066) |
| $n$ | 400 | 0.108 (0.060) | 0.140 (0.098) | 0.150 (0.080) | 0.102 (0.110) | 0.106 (0.098) |
| | 600 | 0.122 (0.096) | 0.120 (0.088) | 0.116 (0.076) | 0.064 (0.090) | 0.098 (0.102) |
| | 800 | 0.096 (0.082) | 0.104 (0.074) | 0.124 (0.104) | 0.118 (0.096) | 0.118 (0.098) |

Table 3. *Empirical powers of $T_n$ and $T_n'$ at nominal level 0.1 when $\omega^2$ changes from 0 to 0.5 and $\sigma^2 = 1$ remains the same. The figures for $T_n'$ are within parentheses*

| $\phi$ | | $0 \rightarrow 0$ | $0 \rightarrow 0.3$ | $0 \rightarrow 0.5$ | $0.3 \rightarrow 0.3$ | $0.3 \rightarrow 0.5$ | $0.5 \rightarrow 0.5$ |
|---|---|---|---|---|---|---|---|
| | 200 | 0.708 (0.364) | 0.826 (0.528) | 0.958 (0.758) | 0.706 (0.392) | 0.820 (0.448) | 0.772 (0.448) |
| $n$ | 400 | 0.918 (0.718) | 0.982 (0.870) | 0.998 (0.942) | 0.930 (0.702) | 0.964 (0.772) | 0.962 (0.778) |
| | 600 | 0.990 (0.874) | 1.00 (0.930) | 1.00 (0.948) | 0.984 (0.848) | 0.990 (0.876) | 0.988 (0.800) |
| | 800 | 0.998 (0.902) | 1.00 (0.922) | 1.00 (0.936) | 1.00 (0.868) | 0.996 (0.872) | 0.994 (0.834) |

Table 4. *Empirical powers of $T_n$ and $T_n'$ at nominal level 0.1 when $\omega^2$ changes from 0 to 0.5 and $\sigma^2$ changes from 1 to 2. The figures for $T_n'$ are within parentheses*

| $\phi$ | | $0 \rightarrow 0$ | $0 \rightarrow 0.3$ | $0 \rightarrow 0.5$ | $0.3 \rightarrow 0.3$ | $0.3 \rightarrow 0.5$ | $0.5 \rightarrow 0.5$ |
|---|---|---|---|---|---|---|---|
| | 200 | 0.994 (0.892) | 0.996 (0.940) | 1.00 (0.952) | 0.994 (0.898) | 0.998 (0.928) | 0.994 (0.868) |
| $n$ | 400 | 1.00 (0.968) | 1.00 (0.960) | 1.00 (0.956) | 1.00 (0.940) | 0.998 (0.922) | 0.996 (0.888) |
| | 600 | 1.00 (0.976) | 1.00 (0.958) | 1.00 (0.960) | 1.00 (0.934) | 1.00 (0.914) | 0.998 (0.876) |
| | 800 | 1.00 (0.968) | 1.00 (0.956) | 1.00 (0.928) | 1.00 (0.952) | 0.998 (0.912) | 1.00 (0.890) |

Table 5. *Empirical powers of $T_n$ and $T_n'$ at nominal level 0.1 when $\omega^2$ changes from 0.1 to 0.5 and $\sigma^2 = 1$ remains the same. The figures for $T_n'$ are within parentheses*

| $\phi$ | | $0 \rightarrow 0$ | $0 \rightarrow 0.3$ | $0 \rightarrow 0.5$ | $0.3 \rightarrow 0.3$ | $0.3 \rightarrow 0.5$ |
|---|---|---|---|---|---|---|
| | 200 | 0.544 (0.254) | 0.676 (0.406) | 0.878 (0.710) | 0.564 (0.270) | 0.688 (0.344) |
| $n$ | 400 | 0.770 (0.550) | 0.926 (0.824) | 0.992 (0.928) | 0.746 (0.516) | 0.870 (0.686) |
| | 600 | 0.856 (0.712) | 0.962 (0.890) | 1.00 (0.956) | 0.888 (0.750) | 0.950 (0.824) |
| | 800 | 0.914 (0.838) | 0.994 (0.932) | 1.00 (0.952) | 0.960 (0.840) | 0.980 (0.878) |

Table 6. *Empirical powers of $T_n$ and $T_n'$ at nominal level 0.1 when $\omega^2$ changes from 0.1 to 0.5 and $\sigma^2$ changes from 1 to 2. The figures for $T_n'$ are within parentheses*

| $\phi$ | | $0 \rightarrow 0$ | $0 \rightarrow 0.3$ | $0 \rightarrow 0.5$ | $0.3 \rightarrow 0.3$ | $0.3 \rightarrow 0.5$ |
|---|---|---|---|---|---|---|
| | 200 | 0.978 (0.860) | 0.990 (0.908) | 0.998 (0.948) | 0.978 (0.842) | 0.972 (0.848) |
| $n$ | 400 | 0.998 (0.970) | 0.996 (0.974) | 0.998 (0.960) | 1.00 (0.952) | 0.998 (0.950) |
| | 600 | 1.00 (0.972) | 1.00 (0.966) | 0.998 (0.960) | 1.00 (0.942) | 0.994 (0.920) |
| | 800 | 1.00 (0.984) | 1.00 (0.962) | 0.998 (0.954) | 1.00 (0.926) | 0.998 (0.936) |

All these results indicate that $T_n$ performs adequately for the parameter change test in the RCA model. In fact, $T_n'$ also produces reasonably good powers, and neither of the two test statistics completely outperforms the other. However, one has to recall that Nyblom's test is no longer applicable if the underlying distribution is completely unknown. In actual practice, the most interesting task is to test for a change in $\omega^2$, especially a change from zero to non-zero $\omega^2$, as ignoring random effects can lead to a false conclusion, for instance, in an interval estimation for the regression parameter.

Table 7. *Empirical sizes and powers of $\Lambda_n$ and $Z_n$ at nominal level 0.1. The figures for $Z_n$ are within parentheses*

| $\phi$ | | Size | | | | Power | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0.1 | 0.3 | 0.5 | 0.7 | 0.1 | 0.3 | 0.5 | 0.7 |
| | 200 | 0.062 | 0.066 | 0.084 | 0.136 | 0.076 | 0.386 | 0.886 | 0.998 |
| | | (0.066) | (0.096) | (0.144) | (0.156) | (0.066) | (0.158) | (0.326) | (0.604) |
| $n$ | 400 | 0.084 | 0.070 | 0.060 | 0.128 | 0.136 | 0.728 | 1.00 | 1.00 |
| | | (0.078) | (0.106) | (0.152) | (0.196) | (0.098) | (0.258) | (0.622) | (0.938) |
| | 600 | 0.084 | 0.084 | 0.072 | 0.126 | 0.208 | 0.904 | 1.00 | 1.00 |
| | | (0.104) | (0.122) | (0.124) | (0.222) | (0.120) | (0.360) | (0.836) | (0.996) |

Now we deal with the test $\Lambda_n$ in section 4. The $\Lambda_n$ is compared with Picard's test statistic $Z_n$. For empirical sizes, sets of 200, 400, and 600 observations are generated from the AR(1) model with $\phi = 0.1, 0.3, 0.5,$ and 0.7 with $\{\epsilon_t\} \sim N(0, 1)$. Here, we test for a change of $\gamma(0)$ and $\gamma(1)$, viz., $m = 1$. In order to see the power, we consider the alternatives under which the first half of the observations are from the AR(1) model described above, and the other half are from i.i.d $N(0, 1/(1 - \phi^2))$ random variables. Here, we keep $\gamma(0)$ constant under both the null and the alternative hypotheses in order to concentrate on $\gamma(1)$. Note that under the null hypothesis, $\gamma(1)$ is $\phi/(1 - \phi^2)$ for the whole set of observations, and under the alternative, $\gamma(1)$ changes from this value to zero. In this simulation, we utilize $h_n = n^{0.4}$ for calculating $\hat{\Gamma}_{ij}$ and $\hat{\kappa}_4 = n^{-1} \sum_{t=1}^{n} \hat{\epsilon}_t^4 / (n^{-1} \sum_{t=1}^{n} \hat{\epsilon}_t^2)^2 - 3$, where $\hat{\epsilon}_t = x_t - \hat{\phi} x_{t-1}$ and $\hat{\phi}$ is the least-squares estimator. The critical value for $\Lambda_n$ at $\alpha = 0.1$ is 2.054 (see Table 1). All the figures in Table 7 are obtained from 500 repetitions.

Table 7 shows the empirical sizes and powers of $\Lambda_n$ and $Z_n$ (the figures within parentheses are for $Z_n$). From Table 7, we can see that the sizes of $\Lambda_n$ are less than 0.1 in most cases, which ensures no severe size distortions. On the other hand, Table 7 shows that $Z_n$ has severe size distortions at $\phi = 0.7$. The power in both cases increases to 1 as either $\phi$ or $n$ increases. It appears that $\Lambda_n$ tends to produce much better powers than $Z_n$. Overall, the result indicates that our cusum test works more appropriately.

## 6. Concluding remarks

In this paper, we proposed a cusum test for parameter change in time series models, and provided a sufficient condition under which the test statistic converges in distribution to the sup of the sum of independent standard Brownian bridges. In section 5, we have seen the simulation results on the cusum test for the models that we discussed in sections 3 and 4. The RCA model and the infinite order moving average model are important models in time series analysis, and the change point test for the regression parameter and autocovariance functions are of much interest to practitioners. As a matter of fact, this paper has been motivated by the fact that the cusum test has not been established in general situations. Although we do not pursue it here, the application of our method is not only restricted to the models which we discussed in our paper, but covers a broad class of statistical models. For instance, one may perform a test for other time series models such as threshold models and ARCH models, or perform a robust test by employing robust estimators if one has concerns about outliers in the data. In this paper, we did not discuss all these cases. However, we believe that our test provides a functional tool to test for a change under a variety of circumstances.

## Acknowledgements

## References

Bai, J. (1994). Weak convergence of the sequential empirical processes of residuals in ARMA models. *Ann. Statist.* **22**, 2051–2061.

Billingsley, P. (1968). *Convergence of probability measures.* Wiley, New York.

Brown, R. L., Durbin, J. & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *J. Roy. Statist. Soc. Ser. B* **37**, 149–163.

Csörgő, M. & Horváth, L. (1988). Nonparametric methods for change point problems. In *Handbook of statistics*, Vol. **7** (eds P. R. Krishnaiah & C. P. Rao), 403–425. Elsevier, New York.

Csörgő, M. & Horváth, L. (1997). *Limit theorems in change-point analysis.* Wiley, New York.

Durbin, J. (1973). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.* **1**, 279–290.

Feigin, P. D. & Tweedie, R. L. (1985). Random coefficient autoregressive processes: a Markov chain analysis of stationarity and finiteness of moments. *J. Time Ser. Anal.* **6**, 1–14.

Gaenssler, P. & Haeusler, E. (1986). On martingale central limit theory. In *Dependence in probability and statistics* (eds E. Eberlein & M. S. Taqqu), 303–334, Birkhäuser, Boston, MA.

Hinkley, D. V. (1971). Inference about the change-point from cumulative sum tests. *Biometrika* **58**, 509–523.

Inclán, C. & Tiao, G. C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variances. *J. Amer. Statist. Assoc.* **89**, 913–923.

Kim, S., Cho, S. & Lee, S. (2000). On the cusum test for parameter changes in GARCH(1,1) models. *Commun. Statist. Theory Methods* **29**, 445–462.

Krämer, W., Ploberger, W. & Alt, R. (1988). Testing for structural change in dynamic models. *Econometrica* **56**, 1355–1369.

Krishnaiah, P. R. & Miao, B. Q. (1988). Review about estimation of change points. In *Handbook of statistics*, Vol. **7** (eds P. R. Krishnaiah & C. P. Rao), 375–402. Elsevier, New York.

Lee, S. (1996). Sequential estimation for the autocorrelations of linear processes. *Ann. Statist.* **24**, 2233–2249.

Lee, S. & Park, S. (2001). The cusum of squares test for scale changes in infinite order moving average processes. *Scand. J. Statist.* **28**, 625–644.

Lee, S. & Wei, C. Z. (1999). On residual empirical process of stochastic regression models with applications to time series. *Ann. Statist.* **27**, 237–261.

Nicholls, D. F. & Quinn, B. G. (1982). *Random coefficient autoregressive models: an introduction.* Springer-Verlag, New York.

Nyblom, J. (1989). Testing for the constancy of parameters over time. *J. Amer. Statist. Assoc.* **84**, 223–230.

Page, E. S. (1955). A test for change in a parameter occurring at an unknown point. *Biometrika* **42**, 523–527.

Phillips, P. C. & Solo, V. (1992). Asymptotics for linear processes. *Ann. Statist.* **20**, 971–1001.

Picard, D. (1985). Testing and estimating change-points in time series. *Adv. Appl. Probab.* **17**, 841–867.

Quinn, B. G. & Nicholls, D. F. (1981). The estimation of random coefficient autoregressive models. II. *J. Time Ser. Anal.* **2**, 185–203.

Stout, W. F. (1970). The Hartman–Wintner law of the iterated logarithm for martingales. *Ann. Math. Statist.* **41**, 2158–2160.

Tang, S. M. & MacNeil, I. B. (1993). The effect of serial correlation on tests for parameter change at unknown time. *Ann. Statist.* **21**, 552–575.

Tong, H. (1990). *Non-linear time series: a dynamical system approach.* Clarendon, Oxford.

Wichern, D. W., Miller, R. B. & Hsu, D. A. (1976). Changes of variance in first-order autoregressive time series models – with an application. *Appl. Statist.* **25**, 248–256.

Zacks, S. (1983). Survey of classical and Bayesian approaches to the change-point problem: fixed sample and sequential procedures of testing and estimation. In *Recent advances in statistics* (eds M. H. Rivzi, J. Rustagi & D. Sigmund), 245–269, Academic Press, New York.

Sangyeol Lee, Department of Statistics, Seoul National University, Seoul, Korea 151-742.
E-mail: sylee@stats.snu.ac.kr.